DEPARTMENT OF DEFENSE BLOGGERS ROUNDTABLE WITH RANDY HILL, EXECUTIVE DIRECTOR, INSTITUTE FOR CREATIVE TECHNOLOGIES, UNIVERSITY OF SOUTHERN CALIFORNIA VIA TELECONFERENCE TIME: 2:00 P.M. EDT DATE: TUESDAY, APRIL 7, 2009

(Note: Please refer to www.dod.mil for more information.)

LINDY KYZER (Army Public Affairs): Okay. Well, I have 2:00 P.M. now, so we'll go ahead and get started. Again, this is Lindy Kyzer with Army Public Affairs. Thank you so much for being here with us.

We're very pleased to have Dr. Randy Hill, who's the executive director for the Institute for Creative Technologies at the University of Southern California. He's here to talk a little bit about what they do, the Institute for Creative Technologies, specifically their work with immersive technology, as well as taking your questions.

So with that, I will go ahead and turn it over to the expert. Dr. Hill?

MR. HILL: Thanks very much, Lindy.   It's a real pleasure to be here today. And I'm just going to give you a little bit of an introduction to the institute, and then I'd be very happy to answer questions.

ICT is a part of the University of Southern California, but we are what is known as a university-affiliated research center or UARC that was established under the sponsorship of the United States Army 10 years ago this summer. And we are in really a strategic relationship with the Army and DOD to begin to address the needs of the Army for training using the technologies that we're developing here that, as Lindy said, are really all about immersion.

The institute was created 10 years ago as a way of establishing a relationship between the entertainment industry, which includes Hollywood and the game industry, university research in virtual reality and the Army. So it's a partnership among those three.

The idea here was that they felt and we feel that we could create -- using Hollywood magic and game-industry magic, we could create learning experiences that engage the mind, the body and the emotions. And we believe that this will result in learning that is deep, learning that sticks.

And the key is really about what we think of as engagement. Engagement is kind of a nebulous term, but I think it means -- it's the ability to catch one's attention.

It's the ability to really draw them in and -- and immerse them in a way that, you know, they just don't want to let go of that experience.

It's -- we define virtual reality -- and this is where the technology part comes in. We broadly define that as any form of interactive digital media that can be hosted on platforms ranging from mobile devices to warehouses. I mean, it's a broad spectrum of experiences that one could have with media.

And the focus of the institute has really been to develop those technologies and to work with the entertainment industry in those -- in the creation of applications that use their magic of storytelling and of interactive experience. But our focus has been on what we think of as the human dimension. We define the human dimension as really being able to look at both the -- all the moral, the physical and cognitive components of soldier development and leadership and organizational leadership.

We have developed applications using these technologies that have addressed issues of recruiting, through interactive characters. We've addressed issues of what we think of as positive training, going beyond just teaching someone how to operate a vehicle or a weapon or something like that. We have focused more on the application of their knowledge in a specific context.

And so you can imagine, for instance, a field artilleryman. You know, they've been trained on how to call for fire. And what we do is we create an environment. Like you can imagine walking into an apartment room that looks like it's in Baghdad, and looking out a synthetic window in that room, that physical room, and doing their call for fire in an urban environment. Think of this as cognitive training because it's training them how to make decisions under stress, and how to perform their tasks and the skills that they have -- use the skills that they have under stress.

We've also -- there's been a huge emphasis here, you know, in terms of this human dimension, on the development of cultural awareness and developing applications that will enable soldiers and leaders to be more effective in cultural situations, through negotiation or just even in meetings that they have or conduct with people in other cultures.

And another area where we've -- we think of as the human dimension is in -- you know, it's really -- it's not the recruiting;   it goes beyond recruiting and training. But really, it has to do with what soldiers face when they come back from a deployment.

And specifically, we've developed a virtual-reality system for treating post-traumatic stress disorder that is now in clinical trials in about 30 different sites around the country.

But I think, overarching, the applications have focused on leadership development, and it's the ability to interact with other human beings in these interpersonal relationships, to motivate them and to understand their problems and interact with them effectively.

Now, there are several key elements to our research program here that make us unique and that I think are -- you know, that are -- you'll see kind of in the DNA of all the applications that we've developed so far.

The six elements are: immersion, which is the creation of mixed reality or digital environments that are meant to engage all of the human senses. We think of a mixed-reality environment as combining physical space, like a room with -- that has physical objects in it, like chairs or tables and

doors -- but combining that with digital objects, which may be projected onto screens in the room or on the walls of the room. And those digital objects may make you, you know, be able to look out a digital window or see a character in a room that you want to have a conversation with.

A second key element is what we think of as virtual humans. This research is focusing on the creation of digital characters with whom you can have a meaningful interaction. And when I think -- when I say a meaningful interaction, I mean it goes beyond what you see in many of the games today, which, you know, are -- tend to be first-person shooters. It's not always the case, but that -- the -- you know, the primary interaction is shooting another individual.

We are focusing on the technologies that will allow you to have a conversation with this character, a conversation where it's not only the spoken word but it focuses on the nonverbal part of the communication, which is the body language, the facial expressions, the gestures -- so that not only can that virtual human perceive your body language, but it's also projecting a part of what it wants to communicate through its posture, facial expressions, and through the spoken word.

The third area of research is in the area of graphics. We want the objects, environments and virtual humans that we create to look very real. And our graphics lab has been very successful at pushing the boundaries of photorealism to the point that a lot of their work is being now used in Hollywood blockbusters. Like "Spider-Man," "Spider- Man 2," "Hancock" and "The Mysterious (sic) Case of Benjamin Button" are examples of movies that have used their technology in which they've received credit.

And the other three elements are games and simulation, which is one where the games and simulations provide a platform for delivery training, where we build on infrastructure that is available commercially and we enhance it with our virtual humans and other content that we create. And we also leverage game industry techniques for designing interactive experiences.

A fifth element is what we think of as narrative and storytelling. And this is really building on an art that has been around for thousands of years, which is -- you know, how you communicate knowledge from one human to another is through a story. And so what you'll see throughout our projects here is this element of storytelling which is a very -- you know, if you think about immersion and engagement and you think about the -- what a book can do, you know, that helps you to just block out -- as you're reading a good novel, it may be able to block out everything else that's happening around you, because you're attached to what's going on with the character and you care about that character and the circumstances they face. We weave that into our training systems and to the films that we've made.

And finally, our key element is the learning sciences. And this is that we're not interested just in creating really cool games or simulations or movies that are entertaining or engaging, but we want them to be effective learning experiences. And so we're taking what is known about instructional design, again, and weaving that into the applications we develop here with the technologies that we're developing; as well as developing intelligent tutoring systems that provide feedback to the students.

Now, so far, we have developed a number of different applications that are being -- you know, that are using the technologies that have been developed

at ICT, and I'd say -- I'd estimate now that we've trained well over 10,000 soldiers using these various applications.

But probably the best thing for me to do at this point is open it up to questions, and I can tell you more based on what your interests are. MS. KYZER: Wow. I certainly think you can. That was a very thorough introduction. I appreciate that. I knew you were involved in a lot, but I learned a lot more in the past couple minutes.

Again, Sean Gallagher with Defense Systems, did you have a question? I'll let you start us off.

Q   Okay. Wanted to dive down a little bit into the virtual person -- the virtual human, I guess you called it.

MR. HILL: Yes.

Q   What -- can you talk a little bit about the technologies that are being developed there for -- in terms of real-language interpretation and the other characteristics of communication with a simulated person?

MR. HILL: Sure. So diving down, there are several key components. The first is the perception component of the character. And this combines not only speech recognition, which -- you know, we know that there are a lot of commercial, off-the-shelf applications of speech recognition out there that are pretty good, but we are working with researchers at USC to advance the state of the art in speech recognition, including doing research on being able to recognize the emotion -- the specific emotions that may be indicated in the prosody of the speech. We also -- and this is the work of Shri Narayanan at USC.

We are also working in the area of computer vision, to be able to track the expressions on the face and the motions of the head that the human is portraying to the virtual human in this interaction. And this is the work of Louis-Philippe Morency.

And this is all, again, part of the perception module that is going to give input to the virtual human in terms of what is being stated or communicated to the -- to the virtual human.

This -- what comes out from the speech recognition and from this -- these -- this computer vision is a set -- an annotated text that then goes out to a -- basically, a language module that is going to try to understand, you know, and parse what has been said -- parse the text strings and the interpretation of the facial expressions.

And this language-understanding module then is fed into what we think of as a cognition module that is -- it's tracking the dialogue, what's been said before, how it relates to what's been said before. And it then -- and it also gauges an emotional response for that virtual human and plans a response that's then sent out to a generation module that is both a -- language generation and generates gestures that the virtual human will show to the -- to the human participant with whom it's interacting.

And this all comes out in the form of speech and gesture that's rendered on a -- you know, on the screen.

That pipeline has to work very fast, as you can imagine. And it's --
you know, because you -- to have a reasonable response, I mean, I think within
50 to 100 milliseconds you've got to have something coming out of that pipeline
once the human has said what they're going to say and you're waiting for a
response.

Q   Okay. So is there a -- are you looking at the provisions for
doing that in duplex? Because, I mean, most people don't have conversations
where there isn't some sort of interruption.

MR. HILL: Yeah. I mean, I think that is one of the research
challenges for us, is the -- is, you know, being able to, you know, exactly deal
with the issue of overlap, because you want them to be processing and
understanding, you know, before the sentence ends. And especially if they're
saying things -- you know, multiple sentences. And so this is an area that's
ongoing in the research by the natural language group that's led by David Traum
here at ICT and Ed Hovy over at the institute -- Information Sciences Institute,
which is part of USC.

MS. KYZER: Okay. And Sean, did you have any other questions?

Q   Yeah, I was -- the -- so in terms of capturing the gestures you
talked about, is that being done with -- are you doing -- are you using motion-
capture technology to bring in gestures or are you looking at some other way of
doing that through animation?

MR. HILL: Yeah. So we've used -- we've used the mo-cap technique to
generate gestures in the past. And what you do it you end up with a sequence of
gestures that you capture that you can then play. But what we try to do is
refine that, much in the same way speech synthesizers work, where you -- you
know, they broke it down into these little phonemes that are generated and
spliced together. We want to be able to splice smaller and smaller gesture
components together, based on what the situation is.

And so we have a project here called SmartBody that takes a behavioral
mark-up language that is generated by the cognition component when it's
generating the response to the human, and the -- that BML goes into the
SmartBody that then generates the gesture, but it comes from a library of
gestures, as opposed to just a long mo-cap stream. Q   Okay.

MS. KYZER: Okay. And I think David Axe may have just joined us.
David, are you on the line?

Q   Yeah, I did. Hi, thanks.

MS. KYZER: Do you want to jump in with a question, or do you have a
topic that maybe you'd like Dr. Hill to open up with or explain a little bit
more?

Q   No. I'm just -- by jumping in like this, I really have no clue
what's going on. So I'm sorry. (Inaudible.)

MS. KYZER: You didn't do your pre-research? Goodness sakes.

Q   (Laughs.)

MS. KYZER: Okay, then. Sean, did you have any other questions?

Q   Wanted to talk a little bit about the computing tech -- the computing back end. You talked about how it needed to be quick response.

It sounds like, from what you were describing with some of the training environments you're trying to set -- they're going to have to be fairly distributed as well, as you're looking at projecting into different rooms. Is that correct?

MR. HILL: Yeah. And kind of one of our ultimate mixed reality environments would be, like I said, a warehouse or multi-room facility, where you might have multiple characters in those rooms and different scenes in those rooms.

Q   So what kind of -- to get that kind of response time and also be able to have multiple instantiations running within that large simulation environment, what kind of computing resources do you need on back end for that?

MR. HILL: You know, what's been amazing to me over the last 10 years, as I've watched this project, is how the computing power has caught up with the needs of things like the Virtual Human Project. We started off running this on a large Reality Monster, which is, as -- you know, it was a Silicon Graphics machine, huge multiprocessor thing, costs a lot of money.

We're now running the same thing -- the same software on a laptop. So the -- and this is a combination of the increase in computing power but also getting smarter about how we run the algorithms.

And so basically, you know, with, you know, higher-end PCs, we can run the software. We don't need any special hardware to do this, and that's just a matter of populating the environment with enough of these for the various rooms. So it doesn't require anything really special.

Q   Okay.

Q   Okay. Hi. This is David.

MR. HILL: Hi, David.

MS. KYZER: Go ahead.

Q   Great. You may have covered this. I'm sorry that I'm late. But -- so we're looking at an immersive training environment that's essentially like a video game you can live in. Right? MR. HILL: Yes.

Q   Great. So what I want to know is, how -- where do you get the -- it's not the magic of the programming and the games, you know, the look of the thing and the hardware that makes it all possible. What I want to know is, where do you get the information to program the behaviors of the players in this game? In other words, how do you -- who teaches these simulated characters to behave like human beings? And where do they -- and how do they know how to teach them? How do they know what to teach them to do?

MR. HILL: Yeah. So it's actually a -- it's a -- this place is very interdisciplinary, and I think that's one of the things that makes us special, is, it's a combination of computer scientists who are also cognitive scientists, and they're working off of theories of emotion and theories of expression, but

then working with subject matter experts for -- depending on what the domain is, if it's -- say if it's -- we've created a character named "Sergeant Star," for instance.

And "Sergeant Star" is an interactive character that the recruiting command is using right now to -- it goes around in the Army interactive van and goes to different facilities where they're going to interact with the public, and he answers questions about the Army.

And he was modeled after a real-life sergeant who's part of the recruiting command, who came in, was scanned in our light stage here at ICT and modeled into a digital character. And it's made to look and act like that particular person. So not only are you working with -- not only do you have computer scientists, but you also have -- we're working from real humans to create the models of gesture and speech pattern. But then we also work with experts, let's say if it's a different culture, on -- you know, if it's, say you're doing a scenario in Iraq or Afghanistan. You would want to work with people who were -- are from that culture, to tell you about -- so that you can come to an understanding of what those patterns of behavior are, so that you can implement them.

And finally, you know, I think this is where Hollywood comes in; is we work with people who are writers, Hollywood storywriters who understand -- you know, they understand human nature and they understand how -- what looks right, feels real, and they add that element to it, too. So that what you come out in the end with is, you know, something that's going to be, you know, hopefully, theoretically realistic as well as seeming real to the people who interact with it.

Q   It just seems to me that the real challenge in creating an environment that's, you know, lifelike but also a useful training tool is that you essentially have to think like -- I'm reluctant to say this, but like an "other," like the "other." And that "other" may even include an enemy. And there's some serious limitations in doing -- sorry?

MR. HILL: I'm agreeing with you.

Q   Oh, okay. There's some serious limitations in trying to do that, because that's sort of fundamentally our whole problem. Isn't that where conflict comes from? So I mean, it seems to me that it's pretty much an impossible challenge. And how close can you -- do we think we can get to simulating the world that -- you know, that confuses us so much as is? (Chuckles.)

MR. HILL: Yeah. Well, I think that you've hit on a key point, and that is, how do you represent people that you may not understand? And the -- what we do, again, is a combination of theoretical work, using tools like we have, for instance, a tool called PsychSim that was developed here based on a -- it's based on the theory of mind, and it's a way of doing social simulations. We have multiple people in the environment, and it's a way of representing what they know. You know, it's what they know about themselves, what they know about others and what they think others think of them, so that if we have a simulation that's running, let's say, in an urban environment, we want to know, well, what's the impact of a particular action going to be on this part of the city or this particular people group, we can simulate that impact on -- based on what we know about those -- you know, that particular group, whether it's religious or political or just cultural group.

Now, part of the way we extract the information that goes into that model is through what we call a cognitive task analysis. And this is a -- this is a pretty well-known technique that's been developed both by Gary Klein and Richard Clark here at USC in the School of Education for finding out about how experts perform their tasks in an environment. And it's also a technique that can be used to extract -- really, to extract tacit knowledge about, you know, what works and what doesn't work at a particular environment, and has been used successfully in extracting knowledge from experts in surgery as well as from, you know, experts in, say, counterinsurgency.

Now, they -- and so what we do is we work with people who've been very successful in operating in a particular, let's say, environment, whether it's in Baghdad or wherever it is and extract that knowledge from them through this technique called cognitive task analysis. And then we try to encode that into the training experience.

So it's not -- again, it's -- you know, it is your best estimate of what is real. And because you're dealing with humans, it's never going to be 100 percent accurate, but you're trying to get it into the zone that represents what could possibly happen.

And, you know, second way we can do this is through looking at stories or accounts of what have happened (sic) in theater. And we have had -- we have some research going on here by Andrew Gordon that is a way -- has a way of extracting stories from large sources of text that it -- and not only identifies the story but indexes it by activity type so that as you're doing your research for a particular application you can get a lot of examples of what's happened in the past through these stories. And he's done this -- he's tested this over the web and identified literally millions of weblogs, identified stories within those weblogs and, again, indexed them. So that you can then type in a prototype story, which you might want to use for doing your research -- could be about a -- you just type in a vanilla version of a fishing trip, for instance, and it'll extract all the fishing stories that it has in its database.

So it's a rigorous process, and it's a challenge. And I think one of the greatest challenges of the 21st century is to build virtual humans that are going to seem real, and that are going to be real in the sense of representing how real people would interact.

Q   Okay. Thank you.

MR. HILL: Sure.

MS. KYZER: And any other questions out there?

Okay, we're actually right at -- right at 2:30 on my clock, so we're doing pretty well here. We'll go ahead and close out. I wanted to turn it back over to Dr. Hill in case he had any closing remarks or anything that we didn't touch on.

MR. HILL: Okay. I think that -- I just wanted to, first of all, acknowledge the great support we've received from the Army, from Dr. Killian (sp), who's a chief scientist, from Dr. John Parmentola, who's really -- oversees the research that supports our work.

But this -- you know, our goal is to have an impact on soldiers -- and a very positive impact, in the sense of enabling them to be more effective as leaders and to be more effective as decision-makers under stress. And, you know, we're very proud of the fact that we've been fulfilling that mission not only in doing cutting-edge research -- I'd say we have, really, the -- some of the top -- one of the top labs in the world when it comes to virtual humans and the graphics -- but we have also been developing applications that are useful, including a couple have -- that have recently won awards within the Army. One is the ELECT BiLAT system for teaching negotiation and meeting skills in a cross-cultural context; and another one called the distribution management cognitive trainer that is used to train logistics people.   And I'm very proud of the fact that this PTSD system is now being used in over 30 sites across the country, including Walter Reed and Madigan Army Medical Center.

And we believe it's going to have a huge impact on the Army and the Defense Department's ability to treat post-traumatic stress disorder. Still in clinical trials, but it's shown a lot of promise, and we think that it's going to have a big impact.

So we are excited to be part of this community and to be having an impact, and we're very thankful for the support we've received from the Department of Defense and the Army.

MS. KYZER: Excellent. Well, thank you so much, sir. Having been able to see a lot of the technology firsthand, it's definitely very cool in person. And I'll try to get some images to the folks who are on the line here, because they're definitely working some very cool stuff.

So thank you so much for your time, sir. This concludes the roundtable.

MR. HILL: All right. Thank you.

END.